REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)* 01/31/2012	2. REPORT TYPE Final Report	3. DATES COVERED *(From - To)* 04/15/2009 - 11/30/2012

4. TITLE AND SUBTITLE		5a. CONTRACT NUMBER
OPTIMIZATION ALGORITHMS AND EQUILIBRIUM ANALYSIS FOR DYNAMIC RESOURCE ALLOCATION		
		5b. GRANT NUMBER FA9550-09-1-0306
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Yinyu Ye		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Stanford University Huang 308, Stanford, CA 94305	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Donald W. Hearn, Program Manager, Optimization and Discrete Mathematics Mathematics, Information and Life Sciences, Air Force Office of Scientific Research	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-OSR-VA-TR-2012-0533

12. DISTRIBUTION / AVAILABILITY STATEMENT

For public-A

13. SUPPLEMENTARY NOTES

14. ABSTRACT

We consider optimization and equilibrium models and algorithms for dynamic resource allocation. The most important accomplishment would be the paper: "The Simplex and Policy-Iteration Methods are Strongly Polynomial for the Markov Decision Problem with a Fixed Discount Rate," where I proved that the classic policy-iteration method (Howard 1960), including the Simplex method (Dantzig 1947) with the most-negative-reduced-cost pivoting rule, is a strongly polynomial-time exact algorithm for solving the Markov decision problem (MDP) exactly with any fixed discount factor. Markov decision process (e.g., Shapley, 1953) is arguably one of the most widely used decision models and methodologies in practice, as a celebrated example to showcase the power of optimization to help making sensible decisions in a complex system and stochastic environment. And the two methods are the most used methods but their theoretical complexities were open before this result.
We also explore an online resource allocation in a set of research publications, including prediction market, Internet auction, and etc.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE		7	19b. TELEPHONE NUMBER *(include area code)*

Final Report of
OPTIMIZATION ALGORITHMS AND EQUILIBRIUM ANALYSIS FOR DYNAMIC
RESOURCE ALLOCATION, AFOSR Grant FA9550-09-1-0306
By Yinyu Ye

Abstract: We consider optimization and equilibrium models and algorithms for dynamic resource allocation. The most important accomplishment of the project would be the paper: "The Simplex and Policy-Iteration Methods are Strongly Polynomial for the Markov Decision Problem with a Fixed Discount Rate," where I proved that the classic policy-iteration method (Howard 1960), including the Simplex method (Dantzig 1947) with the most-negative-reduced-cost pivoting rule, is a strongly polynomial-time exact algorithm for solving the Markov decision problem (MDP) exactly with any fixed discount factor. Markov decision process (e.g., Shapley, 1953) is arguably one of the most widely used decision models/methodologies in practice, as a celebrated example to showcase the power of optimization to help making sensible decisions in a complex system and stochastic environment. And the two methods are the most used methods in real world applications but their theoretical complexities were open before our result. We also explore an online and dynamic resource allocation and mechanism design in a set of research publications, including prediction market, Internet auction, spectrum allocation/trading model, sensor network localization, and etc, where demands for resources arrive sequentially and a decision/trade has to make as soon as a (pair) demand order(s) arrives. We develop online algorithms/decision rules, similar to online routing and online combinatorial auctions, for general dynamic resource to achieve near-optimal social utility value and/or resource utilization.

We outline major accomplishments that have been made from the project.

Markov Decision Process

[1] Ye, "The Simplex and Policy-Iteration Methods are Strongly Polynomial for the Markov Decision Problem with a Fixed Discount Rate," Mathematics of Operations Research, 36:4 (2011) 593-603.

Markov decision process (e.g., Shapley, 1953) is arguably one of the most widely used dynamic decision models/methodologies in practice. It has been an integrated part of virtually any textbook on Operations Research, as a celebrated example to showcase the power of optimization to help making sensible decisions in a complex system and stochastic environment. Although many heuristic and linear programming methods have been proposed and well studied, it had been a long standing open problem to find a strongly polynomial time algorithm (independent of the problem data) for solving the Markov decision problem (MDP). Ye (Ye, "A new complexity result on solving the Markov decision problem," Mathematics of Operations Research, 30:3 (2005) 733-749) resolved this open problem partially first in There, he developed a novel combinatorial interior point algorithm, and proved a strongly polynomial-time bound for solving the MDP problem exactly when the discount factor is fixed. This was the *first* strongly polynomial-time algorithm for MDP even with fixed discount factors. The only

previously known result is a strongly polynomial-time algorithm (Papadimitriou, C. H., J. N. Tsitsiklis, 1987) for the deterministic MDP (that is reduced to a minimum-cycle network flow problem), which is based on Karp's minimum-cycle network flow algorithm.

More impressively, very recently Ye, in [1], proved that the classic policy-iteration method (Howard 1960), including the Simplex method (Dantzig 1947) with the most-negative-reduced-cost pivoting rule, is also a strongly polynomial-time exact algorithm for solving the Markov decision problem (MDP) exactly with any fixed discount factor. Furthermore, the computational complexity of the policy-iteration method (including the Simplex method) is better than that of the interior-point algorithm, which matches its superior practical performance. The result is surprising because the simplex method with the same pivoting rule was shown to be exponential for solving a general linear programming problem (Klee and Minty, 1972), the simplex method with the smallest index pivoting rule was shown to be exponential for solving an MDP regardless of discount rates (Melekopoglou and Condon, 1994), and the policy-iteration method was recently shown to be exponential for solving undiscounted MDPs under the average cost criterion. This is an amazing result, given the fact that these methods exist for over 50 to 60 years, were studied extensively by many excellent researchers, and were popularly used in real-world applications.

In addition, Ye's analyses were adapted by a group computer scientists (Hansen, Miltersen and Zwick, 2011) to show that the policy or strategy iteration method is strongly polynomial for 2-player turn-based stochastic games with discounted zero-sum rewards. This provides the *first* strongly polynomial algorithm for solving these games, resolving a long standing open problem.

Online Optimization and Mechanism Design

[2] Agrawal, Delage, Peters, Wang, and Ye, "A Unified Framework for Dynamic Prediction Market Design," *Operations Research*, 59:3 (2011) 550-568;
[3] Agrawal, Ding, Sebari, and Ye, "Price of Correlations in Stochastic Optimization" to appear in *Operations Research*.

Recently, coinciding with and perhaps driving the increased popularity of prediction markets, several novel pari-mutuel mechanisms have been developed such as the logarithmic market-scoring rule (LMSR), the cost-function formulation of market makers, utility-based markets, and the sequential convex pari-mutuel mechanism (SCPM). In [2], we present a convex optimization framework that unifies these seemingly unrelated models for centrally organizing contingent claims markets. The existing mechanisms can be expressed in our unified framework by varying the choice of a concave value function. We show that this framework is equivalent to a convex risk minimization model for the market maker. This facilitates a better understanding of the risk attitudes adopted by various mechanisms. The unified framework also leads to easy implementation because we can now find the cost function of a market maker in polynomial time by solving a simple convex optimization problem. In addition to unifying and explaining the existing

mechanisms, we use the generalized framework to derive necessary and sufficient conditions for many desirable properties of a prediction market mechanism such as proper scoring, truthful bidding (in a myopic sense), efficient computation, controllable risk measure, and guarantees on the worst-case loss. As a result, we develop the first proper, truthful, risk-controlled, loss-bounded (independent of the number of states) mechanism; none of the previously proposed mechanisms possessed all these properties simultaneously. Thus, our work provides an effective tool for designing new prediction market mechanisms. We also discuss possible applications of our framework to dynamic resource pricing and allocation in general trading markets.

We believe that our framework of [2] for designing dynamic prediction markets has intimate connections to other dynamic trading markets such as online auction of goods, and could lead to interesting results for these markets as well. In general, any dynamic resource allocation and pricing scheme relies crucially on the trade-off between the profit achieved by exploiting the resource now versus the value of saving the goods for the future and exploring the market further. This future value of resources is captured in our framework by the concave value function. Our risk-based formulation also formalized how this value function captures the trade-off between learning the preferences of the traders versus maximizing instant profit via a penalty function. This bears similarities to the classic exploration versus exploitation trade-off for general trading markets. Additionally, our mechanism achieves incentive compatibility using the VCG allocation and pricing scheme popular for online auctions of goods. Further investigation of implications of our results on other trading and auction markets is part of an ongoing research.

When decisions are made in presence of high dimensional stochastic data, handling joint distribution of correlated random variables can present a formidable task, both in terms of sampling and estimation as well as algorithmic complexity. A common heuristic is to estimate only marginal distributions and substitute joint distribution by independent (product) distribution. In [3], we study possible loss incurred on ignoring correlations through a distributionally-robust stochastic programming model, and quantify that loss as Price of Correlations (POC). Using techniques of cost-sharing from game theory, we identify a wide class of problems for which POC has a small upper bound. To our interest, this class will include many stochastic convex programs, uncapacitated facility location, Steiner tree, and submodular functions, suggesting that the intuitive approach of assuming independent distribution approximates the robust model for these stochastic optimization problems. Additionally, we demonstrate hardness of bounding POC via examples of subadditive and supermodular functions that have large POC. We find that our results are also useful for solving many deterministic optimization problems like welfare maximization, k-dimensional matching and transportation problem, under certain conditions.

In summary, [3] proposed an approximation algorithm to solve the DRSP model that simply ignores the correlations and can be implemented efficiently, and introduced a new concept of POC to measure the approximation ratio achieved. We believe the concept of POC is especially attractive because it characterizes the cases when the seemingly

pessimistic worst case joint distribution is close to the more natural independent distribution, in the sense that former can be substituted by the latter. By proving upper an lower bounds on POC for a wide range of problems, our research sheds important insights on when correlations can be ignored in practice. We also show that many deterministic optimization problems that involve matching or partitioning constraints can be formulated as the problem of computing worst case distribution with given marginals. Hence, our results provide approximation algorithms for those as well. Finally, our methodology of bounding POC using cost-sharing schemes is a novel application of these algorithmic game theory techniques and deserves further study.

Computational Game Theory and Market Equilibrium

[4] Zhu, Dang and Ye, "A FPTAS for Computing a Symmetric Leontief Competitive Economy Equilibrium," *Math Programming* 131 (2012) 113–129.
[5] Dang, Zhu and Ye, "An interior-point path-following algorithm for computing a Leontief economy equilibrium," *Computational Optimization and Applications* 50:2 (2011).
[6] Yao, Armbruster, and Ye "Dynamic Spectrum Management with the Competitive Market Model," *IEEE Tran on Signal Processing* 58:4 (2010) 2442-2446.

The Arrow-Debreu competitive market equilibrium problem which was first formulated by Leon Walras in 1874. In this problem everyone in a population of n players has an initial endowment of a divisible good and a utility function for consuming all goods--- their own and others. Every player produce/sells the entire initial endowment and then uses the revenue to buy a bundle of goods such that his or her utility function is maximized. Walras asked whether prices could be set for everyone's good such that this is possible. An answer was given by Arrow and Debreu (1954) who showed that such equilibrium would exist if the utility functions were concave. Their proof was non-constructive and did not offer any algorithm to find such equilibrium prices.
Fisher (1891) was the first to consider an algorithm to compute equilibrium prices for a special case model where players are divided into two sets: producers and consumers. Scarf in 1973 also developed an algorithm to solve general fixed point problems, including the competitive market equilibrium problem. His algorithm, however, was not proved to be polynomial-time.

If the utility functions are linear, Eisenberg and Gale (1959) gave a nonlinear convex optimization setting to formulate the Fisher model and Nenakhov and Primak (1983, also Jain 2004) gave a nonlinear convex optimization setting to formulate the general Arrow-Debreu model. Thus, the Ellipsoid method can be used to solve them approximately in polynomial time where the bound is of order $O(n^8 L)$. Recenty, Ye developed an interior-point algorithm that solves both the Fisher and Arrow-Debreu models exactly, when the utilities are linear, in polynomial time of order $O(n^4 L)$, which is in line with the best complexity bound for linear programming of the same dimension and size. These results motivated many researchers in the past few years look for polynomial-time algorithms for solving more general utility equilibrium problems, such as the Leontief utility, which is a piece-wise linear concave function. However, soon after Ye and his co-authors announced that the computation of the Arrow-Debreu equilibrium with the Leontief

utility is NP-hard, which effectively terminates any hope to develop a polynomial-time algorithm unless P=NP.

On the positive sides, in [4], we consider a linear complementarity problem (LCP) arisen from the Nash and Arrow–Debreu with Leontief utility competitive economy equilibria where the LCP coefficient matrix is symmetric. We prove that the decision problem, to decide whether or not there exists a complementary solution, is NP-complete. Under certain conditions, an LCP solution is guaranteed to exist and we present a fully polynomialtime approximation scheme (FPTAS) for approximating a complementary solution, although the LCP solution set can be non-convex or non-connected. Our method is based on approximating a quadratic social utility optimization problem (QP) and showing that a certain KKT point of the QP problem is an LCP solution. Then, we further show that such a KKT point can be approximated with a new improved running time complexity. We also report computational results which show that the method is highly effective. Applications in competitive market model problems with other utility functions are also presented, including global trading and dynamic spectrum management problems.

Then, in [5], we present an interior-point path-following algorithm for computing a nonsymmetric Leontief economy equilibrium, that is, an exchange market equilibrium with Leontief utility functions, which is known to be in the complexity class of PPAD-complete. We construct a smooth homotopy interior-point path to tackle this system. We prove that there always exists a continuously differentiable path leading to a complementary solution of the nonlinear system and at the same time to a Leontief economy equilibrium associated with the solution. We also report preliminary computational results to show effectiveness of the path-following Newton method.

Dynamic spectrum management (DSM) is a technology to efficiently share the spectrum among the users in a communication system. DSM can be used in the digital subscriber line (DSL) systems to reduce cross-talk interference and improve total system throughput. DSM is also a promising candidate for multiple access in cognitive radio. In DSM, multiple users coexist in a channel, and this causes co-channel interference. The goal of DSM is to manage the power allocations in all the channels to maximize the sum of the data rates of all the users, subject to power constraints. Unfortunately, this problem is non-convex and cannot be solved efficiently in polynomial time. In [6], have shown that dynamic spectrum management (DSM) using the market competitive equilibrium (CE), which sets a price for transmission power on each channel, leads to better system performance in terms of the total data transmission rate (by reducing cross talk), than using the Nash equilibrium (NE). But how to achieve such a CE is an open problem. We show that the CE is the solution of a linear complementarity problem (LCP) and can be computed efficiently. We propose a decentralized tˆatonnement process for adjusting the prices to achieve a CE. We show that under reasonable conditions, any tatonnement process converges to the CE. The conditions are that users of a channel experience the same noise levels and that the cross-talk effects between users are low-rank and weak.

Optimization for Sensor Network

[7] Zhu, So and Ye, "Universal Rigidity and Edge Sparsification for Sensor Network Localization," to appear in *SIAM J. Optimization*, 2012.
[8] Alfakih, Taheri, and Ye, "On stress matrices of (d + 1)-lateration frameworks in general position," to appear in *Math Programming*, 2012.

Owing to their high accuracy and ease of formulation, there has been great interest in applying convex optimization techniques, particularly that of semidefinite programming (SDP) relaxation, to tackle the sensor network localization problem in recent years. However, a drawback of such techniques is that the resulting convex program is often expensive to solve. In order to speed up computation, various edge sparsification heuristics have been proposed, whose aim is to reduce the number of edges in the input graph. Although these heuristics do reduce the size of the convex program and hence making it faster to solve, they are often ad hoc in nature and do not preserve the localization properties of the input. As such, one often has to face a tradeoff between solution accuracy and computational effort. In [7], we propose a novel edge sparsification heuristic that can provably preserve the localization properties of the original input. At the heart of our heuristic is a graph decomposition procedure, which allows us to identify certain sparse *generically universally rigid* subgraphs of the input graph. Our computational results show that the proposed approach can significantly reduce the computational and memory complexities of SDP–based algorithms for solving the sensor network localization problem. Moreover, it compares favorably with existing speedup approaches, both in terms of accuracy and solution time.

Paper [8] is a technical paper that resolved an important mathematical question on the rigidity of a sensor network and structure. Let (G, P) be a bar framework of n vertices in *general position* in R^d, for d ≤ n − 1, where G is a (d + 1)-lateration graph. In [8], we presented a constructive proof that (G, P) admits a positive semidefinite stress matrix with rank (n−d−1). We also prove a similar result for a sensor network, where the graph consists of m(≥ d + 1) anchors.